Design of Automatic Strawberry Harvest Robot Suitable in Complex Environments

Zhuoling Huang*
zhuoling.huang@kcl.ac.uk
Department of Informatics
King's College London, UK

Elizabeth Sklar
esklar@lincoln.ac.uk
University of Lincoln & King's College London, UK

Simon Parsons
sparsons@lincoln.ac.uk
University of Lincoln & King's College London, UK

ABSTRACT

Strawberries are an important cash crop that are grown worldwide. They are also a labour-intensive crop, with harvesting a particularly labour-intensive task because the fruit needs careful handling. This project investigates collaborative human-robot strawberry harvesting, where interacting with a human potentially increases the adaptability of a robot to work in more complex environments. The project mainly concentrates on two aspects of the problem: the identification of the fruit and the picking of the fruit.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; • **Computer systems organization → Robotic control**; • **Computing methodologies →** *Vision for robotics.*

KEYWORDS

agricultural robot; human robot interaction; learning from human

ACM Reference Format:

Zhuoling Huang, Elizabeth Sklar, and Simon Parsons. 2020. Design of Automatic Strawberry Harvest Robot Suitable in Complex Environments. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20 Companion), March 23–26, 2020, Cambridge, United Kingdom.*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3371382.3377443

1 INTRODUCTION

Strawberries are planted widely around the world and are an important cash crop, but strawberry farming has quite a high labour cost. For example, the total labour time of working on strawberry plants in Japan is slightly less than 20,000 hours/hectare, while harvesting is about 5000 hours/hectare [6], and the harvesting labour cost could sometimes reach 45% of the entire labour cost [18]. These costs motivate a strong need for a strawberry harvesting robot. The environments in which strawberries are grown are usually very *complex*: there are many different varieties of strawberry with different shapes, the farm environment includes a range of very different illumination and background conditions, and strawberries grow in such a way that berries are often obscured by leaves and other fruit. Thus, methods tailored to one specific environment may not work in another. As the goal of this research is to design a robot suitable for these complex environments, a potential way to improve performance in automated strawberry harvesting is to develop intelligent human-robot solutions. Two questions therefore are investigated here: (1) How could humans help to improve the performance of robots? (2) How could humans and robots interact with each other more efficiently? To address these questions two preliminary results are described: a pilot study of human-robot interaction for fruit identification, and research into a robot learning specialised movements for harvesting from a human.

Related Work Approaches to automated strawberry harvesting encompass research from multiple communities; work related to object detection and human-robot interaction are highlighted here. Fruit harvest robots are investigated in many different research projects, including but not limited to strawberries [5, 6, 15, 17, 18]. Different methods are introduced to detect the target object, and the mainstream research on fruit detection can be roughly divided into two phases: detecting with traditional computer vision methods based on shape [12, 14] or colour [6, 11], and detecting with machine learning methods [8, 13]. However, when the robot works in different complex environments, none of these methods on their own can be transferred directly without making adjustments to maintain performance standards. To improve the adaptability of the robot, a solution is to include a human in the loop. Learning from Demonstrations (LfD) provides a sensible approach for a robot to generate a policy from human teachers [1, 9, 10].

Preliminary Work Prior to embarking on human-robot approaches as outlined here, we studied autonomous fruit detection. To detect the picking point of a strawberry (the position that a robot hand should grasp to detach the fruit), we designed and tested an automated vision-based system [7]. First, a target mature strawberry is distinguished from the image background using colour segmentation. Then to enable choosing a picking point in the image, a specially designed template was used. This approach works better than previous works in the same area [4, 18] and was tested on our own dataset containing multiple strawberries images taken in real farm environments. This detection method is then applied in our current work together with detectors based on neural networks.

2 CURRENT WORK

When dealing with some high-value crops, the behaviour of a robot is preferred to be more reliable, thus we explore a human-in-the-loop strategy Taking strawberries as an example, the basic process of fruit harvesting could include tasks such as identifying mature strawberries ready to pick, choosing suitable positions for detaching fruit from the plant, controlling the robotic arm to approach the target, and removing the fruit from the plant. In such complex

*Supported by the China Scholarship Council (No.201808060377).

HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom.

ACM ISBN 978-1-4503-7057-8/20/03.
https://doi.org/10.1145/3371382.3377443

tasks, we consider an approach where human collaborators cooperate with a robot for target identification, as well as the training a controller for the robot arm.

Human Involved Detection Pilot Study We apply supervised machine learning to the task of detecting ripe strawberries in images. However, due to limitations of our training dataset, it is hard for the accuracy of detection to reach 100%, especially when a detector is applied to a very different environment from that where it is trained. A possible way to improve the performance is to include a human in the loop. For example, when the robot is working with a new variety of strawberry that is not included in the training dataset, a human user can help the robot by checking the detection result and providing feedback for further learning. As a collaboration task, it is important for the human collaborator to consider the robot's decision to be trustworthy or at least helpful. Thus, the impact of different ways that robot sharing decision making with a human when detecting mature strawberries is studied.

For this study, an indoor simulated strawberry farm was created: 52 high-resolution strawberry images taken on real farms were printed out in full-colour and hung on the walls of the corridor outside our robotics lab. An Asus Xtion camera on a mobile Turtlebot2 was then used to take 30 pictures at 5 pre-defined locations of the simulated strawberry farm to conduct a controlled experiment. Two different robot behaviours (algorithms) for detecting strawberries were compared: one detects targets with pre-trained classifier and displays the result in a binary mode (mature strawberry or not), and the other detects mature strawberries with the same classifier but provides a confidence value for its detection. When the detection result is displayed in a graphical user interface, the user who collaborates with the robot is able to check the robot's detection result, remove the label of false positive results and label false negative results. The basic setup is shown in Figure 1(a).

As a pilot study, 13 volunteers took part in an experiment, each of them completed the detection task two times with the robot using the different two algorithms mentioned above in a random order. Questionnaires are used for collecting subjective data on human perception of working with each robot. A pre-survey and a post-survey were given to the users before and after the experiment respectively. Meanwhile, objective data on robot performance and human working time was also collected and analysed. The results of this pilot study showed that the users preferred the algorithm with confidence value provided, which is more informative compared with the binary algorithm. This result follows an overall trend in human-AI interaction systems: a system that explains its reasoning tends to help increase user satisfaction [3].

Robot Arm Motion Learning from Human We apply Learning from Demonstration (LfD) to the task of moving the robot arm to an appropriate location for picking strawberries. The idea is that learning from human teachers will increase the flexibility of arm motion. Compared with traditional manipulation, learning from human teachers enables a robot to be controlled by non-technical-professional users. And different to the linear movement adopted by existing fruit pickers, the trajectory can be more flexible, increasing the possibility of deploying a robot in complex environments.

For the robot to learn from human teacher(s), the whole process is divided into two parts. First, a library is constructed based on demonstrations provided by a human. Thus, a series of movements

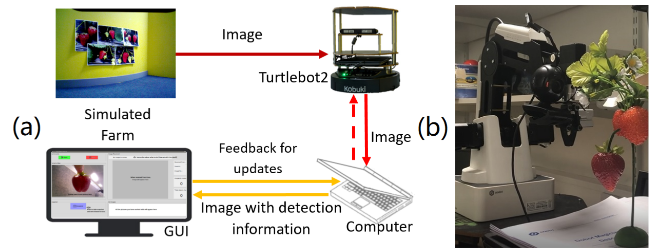

Figure 1: Setup for current work. Image (a) shows the setup for human involved detection pilot study. Image (b) shows the setup for LfD experiments.

was executed, and then the coordinate (X, Y, Z) of the center of the end platform, the rotation angle of the servo joint relative to the coordinate frame, and the 4 joint angle between links of the arm were recorded altogether with time as a whole dataset. The second step is deriving a policy from the library. Since the demonstrations in the dataset are Markovian, the Hierarchical Dirichlet Process Hidden Markov Model (HDP-HMM) [2] is applied to the dataset to cluster the demonstrations into different subsequences. For each group of subsequences, the Dynamic movement primitives (DMPs) method [16] is then used for parameterizing the subsequences. The parameterized trajectories provide a policy as the starting point, with parameters to be explored later using reinforcement learning.

Figure 1(b) shows an experimental setup to control the Dobot Magician arm to reach a target. A webcam is attached to the end effector to take a picture of the environment at a certain frequency. The arm will move around, and when a mature strawberry is detected in the centre of the image and occupies of the image, it is considered to have been reached. This result can be applied to the reward function when using reinforcement learning. To measure the performance of the learned method, the accuracy of robot arm reaching the target is calculated.

3 NEXT STEPS

For human-in-the-loop detection, the first thing to do is to extend our preliminary user study with more volunteers to obtain a larger sample, which will produce more reliable results. In addition, more robot behaviours could be studied, for example, the impact of the rate of false positives and false negatives could be an interesting topic. Finally, the labelling information feedback from users could be applied to reinforcement learning, to help with improving the accuracy of the detectors.

For the robot arm motion learning from human teachers, one possible next step is to improve the performance of humans first, by increasing the dataset with more demonstrations from strawberry harvest professions. In addition, learning based on observations to choose a suitable angle for robots to avoid obstacles when approaching the targets should also be studied.

It is believed that with humans collaborating with robots, the results for fruit harvesting tasks will be better than either working on their own. The robots reduce the repetitive workload of human, and the human helps the robot to adapt in a more complex environment.

REFERENCES

[1] Daniel S. Brown and Scott Niekum. 2018. Machine teaching for inverse reinforcement learning: Algorithms and applications. *arXiv preprint arXiv:1805.07687* (2018).

[2] Emily B. Fox, Erik B. Sudderth, Michael I. Jordan, and Alan S. Willsky. 2008. An HDP-HMM for systems with state persistence. In *Proceedings of the 25th international conference on Machine learning*. ACM, 312–319.

[3] Randy Goebel, Ajay Chander, Katharina Holzinger, Freddy Lecue, Zeynep Akata, Simone Stumpf, Peter Kieseberg, and Andreas Holzinger. 2018. Explainable AI: the new 42?. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE)*. Springer, 295–303.

[4] F. Guo, Q. Cao, and N. Masateru. 2008. Fruit detachment and classification method for strawberry harvesting robot. *International Journal of Advanced Robotic Systems* 5 (2008), 41–48. Issue 1.

[5] Michael Halstead, Christopher McCool, Simon Denman, Tristan Perez, and Clinton Fookes. 2018. Fruit Quantity and Quality Estimation using a Robotic Vision System. *arXiv preprint arXiv:1801.05560* (2018).

[6] Shigehiko Hayashi, Kenta Shigematsu, Satoshi Yamamoto, Ken Kobayashi, Yasushi Kohno, Junzo Kamata, and Mitsutaka Kurita. 2010. Evaluation of a strawberry-harvesting robot in a field test. *Biosystems engineering* 105, 2 (2010), 160–171.

[7] Zhuoling Huang, Sam Wane, and Simon Parsons. 2017. Towards Automated Strawberry Harvesting: Identifying the Picking Point. In *Annual Conference Towards Autonomous Robotic Systems*. Springer, 222–236.

[8] Chris McCool, Tristan Perez, and Ben Upcroft. 2017. Mixtures of lightweight deep convolutional neural networks: applied to agricultural robotics. *IEEE Robotics and Automation Letters* 2, 3 (2017), 1344–1351.

[9] Scott Niekum, Sarah Osentoski, George Konidaris, and Andrew G. Barto. 2012. Learning and generalization of complex tasks from unstructured demonstrations. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5239–5246.

[10] Scott Niekum, Sarah Osentoski, George Konidaris, Sachin Chitta, Bhaskara Marthi, and Andrew G. Barto. 2015. Learning grounded finite-state representations from unstructured demonstrations. *The International Journal of Robotics Research* 34, 2 (2015), 131–157.

[11] YuuIchi. Ohta, Takeo. Kanade, and Toshiyuki. Sakai. 1980. Color information for region segmentation. *Computer graphics and image processing* 13, 3 (1980), 222–241.

[12] Hetal N. Patel, Ramesh K. Jain, and Manjunath V. Joshi. 2012. Automatic segmentation and yield measurement of fruit using shape analysis. *International Journal of Computer Applications* 45, 7 (2012), 19–24.

[13] Steven Puttemans, Yasmin Vanbrabant, Laurent Tits, and Toon Goedemé. 2016. Automated visual fruit detection for harvest estimation and robotic harvesting. In *Image Processing Theory Tools and Applications (IPTA), 2016 6th International Conference*. IEEE, 1–6.

[14] Ribana Roscher, Katja Herzog, Annemarie Kunkel, Anna Kicherer, Reinhard Töpfer, and Wolfgang Förstner. 2014. Automated image analysis framework for high-throughput determination of grapevine berry sizes using conditional random fields. *Computers and Electronics in Agriculture* 100 (2014), 148–158.

[15] Inkyu Sa, Zongyuan Ge, Feras Dayoub, Ben Upcroft, Tristan Perez, and Chris McCool. 2016. Deepfruits: A fruit detection system using deep neural networks. *Sensors* 16, 8 (2016), 1222.

[16] Stefan Schaal. 2006. Dynamic movement primitives-a framework for motor control in humans and humanoid robotics. In *Adaptive motion of animals and machines*. Springer, 261–280.

[17] Eldert J. Van Henten, Jochen Hemming, B. A. J. Van Tuijl, J. G. Kornet, J. Meuleman, J. Bontsema, and E. A. Van Os. 2002. An autonomous robot for harvesting cucumbers in greenhouses. *Autonomous Robots* 13, 3 (2002), 241–258.

[18] Tiezhong Zhang, , Libing Chen, and Song Jian. 2005. Study on strawberry harvesting robot: II. Images based identifications of strawberry barycenter and plucking position. *Journal of China Agricultural University* 10, 1 (2005), 48–51.